Abstract

I describe several techniques for characterizing
molecules based on the shapes of their fields. The minimal
distance between two molecular fields is used as a shape-
based metric, independent of the underlying chemical
structure, and a high-dimensional shape space description of
the molecules is generated. I then show how these attributes
can be used in creating, characterizing, and searching
databases of molecules based on field similarity. In
particular, they allow searches of a database in sublinear
time. Next, I extend the utility of this approach by
describing a way to automatically break molecules into a
series of fragments by using an ellipsoidal Gaussian
decomposition. Not only can these fragments then be analyzed
by the shape metric technique described above, but the
parameters of the decomposition themselves can also be used
to further organize and search databases. The ellipsoidal
method can also be used to describe binding or active sites
on macromolecules, providing a template for searching for
complementary molecules in a database such as I describe.
The most immediate application of these techniques is to
pharmaceutical drug discovery and design.